

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Gray, Patrick W.
- (ii) TITLE OF INVENTION: Chitinase Materials and Methods
- (iii) NUMBER OF SEQUENCES: 15
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 - (B) STREET: 6300 Sears Tower, 233 South Wacker Drive
 - (C) CITY: Chicago
 - (D) STATE: Illinois
 - (E) COUNTRY: United States of America
 - (F) ZIP: 60606-6402
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Rin-Laures, Li-Hsien
 - (B) REGISTRATION NUMBER: 33,547
 - (C) REFERENCE/DOCKET NUMBER: 27866/32960
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 312/474-6300
 - (B) TELEFAX: 312/474-0448
 - (C) TELEX: 25-3856

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1636 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 2..1399
- (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) LOCATION: 65..1399
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

C ATG GTG CGG TCT GTG GCC TGG GCA GGT TTC ATG GTC CTG CTG ATG
Met Val Arg Ser Val Ala Trp Ala Gly Phe Met Val Leu Leu Met
-21 -20 -15 -10

46

ATC CCA TGG GGC TCT GCT GCA AAA CTG GTC TGC TAC TTC ACC AAC TGG

94

Ile Pro Trp Gly Ser Ala Ala Lys Leu Val Cys Tyr Phe Thr Asn Trp	10	
GCC CAG TAC AGA CAG GGG GAG GCT CGC TTC CTG CCC AAG GAC TTG GAC	142	
Ala Gln Tyr Arg Gln Gly Glu Ala Arg Phe Leu Pro Lys Asp Leu Asp	25	
15 20		
CCC AGC CTT TGC ACC CAC CTC ATC TAC GCC TTC GCT GGC ATG ACC AAC	190	
Pro Ser Leu Cys Thr His Leu Ile Tyr Ala Phe Ala Gly Met Thr Asn	40	
30 35		
CAC CAG CTG AGC ACC ACT GAG TGG AAT GAC GAG ACT CTC TAC CAG GAG	238	
His Gln Leu Ser Thr Thr Glu Trp Asn Asp Glu Thr Leu Tyr Gln Glu	55	
45 50		
TTC AAT GGC CTG AAG AAG ATG AAT CCC AAG CTG AAG ACC CTG TTA GCC	286	
Phe Asn Gly Leu Lys Lys Met Asn Pro Lys Leu Lys Thr Leu Leu Ala	70	
60 65		
ATC GGA GGC TGG AAT TTC GGC ACT CAG AAG TTC ACA GAT ATG GTC GCC	334	
Ile Gly Gly Trp Asn Phe Gly Thr Gln Lys Phe Thr Asp Met Val Ala	90	
75 80 85		
ACG GCC AAC AAC CGT CAG ACC TTT GTC AAC TCG GCC ATC AGG TTT CTG	382	
Thr Ala Asn Asn Arg Gln Thr Phe Val Asn Ser Ala Ile Arg Phe Leu	105	
95 100		
CGC AAA TAC AGC TTT GAC GGC CTT GAC CTT GAC TGG GAG TAC CCA GGA	430	
Arg Lys Tyr Ser Phe Asp Gly Leu Asp Leu Asp Trp Glu Tyr Pro Gly	120	
110 115		
AGC CAG GGG AGC CCT GCC GTA GAC AAG GAG CGC TTC ACA ACC CTG GTA	478	
Ser Gln Gly Ser Pro Ala Val Asp Lys Glu Arg Phe Thr Thr Leu Val	135	
125 130		
CAG GAC TTG GCC AAT GCC TTC CAG CAG GAA GCC CAG ACC TCA GGG AAG	526	
Gln Asp Leu Ala Asn Ala Phe Gln Gln Glu Ala Gln Thr Ser Gly Lys	150	
140 145		
GAA CGC CTT CTT CTG AGT GCA GCG GTT CCA GCT GGG CAG ACC TAT GTG	574	
Glu Arg Leu Leu Ser Ala Ala Val Pro Ala Gly Gln Thr Tyr Val	170	
155 160 165		
GAT GCT GGA TAC GAG GTG GAC AAA ATC GCC CAG AAC CTG GAT TTT GTC	622	
Asp Ala Gly Tyr Glu Val Asp Lys Ile Ala Gln Asn Leu Asp Phe Val	185	
175 180		
AAC CTT ATG GCC TAC GAC TTC CAT GGC TCT TGG GAG AAG GTC ACG GGA	670	
Asn Leu Met Ala Tyr Asp Phe His Gly Ser Trp Glu Lys Val Thr Gly	200	
190 195		
CAT AAC AGC CCC CTC TAC AAG AGG CAA GAA GAG AGT GGT GCA GCA GCC	718	
His Asn Ser Pro Leu Tyr Lys Arg Gln Glu Ser Gly Ala Ala Ala	215	
205 210		
AGC CTC AAC GTG GAT GCT GCT GTG CAA CAG TGG CTG CAG AAG GGG ACC	766	
Ser Leu Asn Val Asp Ala Ala Val Gln Gln Trp Leu Gln Lys Gly Thr	230	
220 225		
CCT GCC AGC AAG CTG ATC CTT GGC ATG CCT ACC TAC GGA CGC TCC TTC	814	
Pro Ala Ser Lys Leu Ile Leu Gly Met Pro Thr Tyr Gly Arg Ser Phe	250	
235 240 245		
ACA CTG GCC TCC TCA TCA GAC ACC AGA GTG GGG GCC CCA GCC ACA GGG	862	
Thr Leu Ala Ser Ser Asp Thr Arg Val Gly Ala Pro Ala Thr Gly		

DECODED SEQUENCES

	255	260	265	
TCT GGC ACT CCA GGC CCC TTC ACC AAG GAA GGA GGG ATG CTG GCC TAC Ser Gly Thr Pro Gly Pro Phe Thr Lys Glu Gly Gly Met Leu Ala Tyr 270 275 280				910
TAT GAA GTC TGC TCC TGG AAG GGG GCC ACC AAA CAG AGA ATC CAG GAT Tyr Glu Val Cys Ser Trp Lys Gly Ala Thr Lys Gln Arg Ile Gln Asp 285 290 295				958
CAG AAG GTG CCC TAC ATC TTC CGG GAC AAC CAG TGG GTG GGC TTT GAT Gln Lys Val Pro Tyr Ile Phe Arg Asp Asn Gln Trp Val Gly Phe Asp 300 305 310				1006
GAT GTG GAG AGC TTC AAA ACC AAG GTC AGC TAT CTG AAG CAG AAG GGA Asp Val Glu Ser Phe Lys Thr Lys Val Ser Tyr Leu Lys Gln Lys Gly 315 320 325 330				1054
CTG GGC GGG GCC ATG GTC TGG GCA CTG GAC TTA GAT GAC TTT GCC GGC Leu Gly Ala Met Val Trp Ala Leu Asp Leu Asp Asp Phe Ala Gly 335 340 345				1102
TTC TCC TGC AAC CAG GGC CGA TAC CCC CTC ATC CAG ACG CTA CGG CAG Phe Ser Cys Asn Gln Gly Arg Tyr Pro Leu Ile Gln Thr Leu Arg Gln 350 355 360				1150
GAA CTG AGT CTT CCA TAC TTG CCT TCA GGC ACC CCA GAG CTT GAA GTT Glu Leu Ser Leu Pro Tyr Leu Pro Ser Gly Thr Pro Glu Leu Glu Val 365 370 375				1198
CCA AAA CCA GGT CAG CCC TCT GAA CCT GAG CAT GGC CCC AGC CCT GGA Pro Lys Pro Gly Gln Pro Ser Glu Pro Glu His Gly Pro Ser Pro Gly 380 385 390				1246
CAA GAC ACG TTC TGC CAG GGC AAA GCT GAT GGG CTC TAT CCC AAT CCT Gln Asp Thr Phe Cys Gln Gly Lys Ala Asp Gly Leu Tyr Pro Asn Pro 395 400 405 410				1294
CGG GAA CGG TCC AGC TTC TAC AGC TGT GCA GCG GGG CGG CTG TTC CAG Arg Glu Arg Ser Ser Phe Tyr Ser Cys Ala Ala Gly Arg Leu Phe Gln 415 420 425				1342
CAA AGC TGC CCG ACA GGC CTG GTG TTC AGC AAC TCC TGC AAA TGC TGC Gln Ser Cys Pro Thr Gly Leu Val Phe Ser Asn Ser Cys Lys Cys Cys 430 435 440				1390
ACC TGG AAT TGAGTCGCTA AAGCCCCCTCC AGTCCCAGCT TTGAGGCTGG Thr Trp Asn 445				1439
GCCCAGGATC ACTCTACAGC CTGCCCTCTG GGTTTTCCCT GGGGGCCGCA ATCTGGCTCC				1499
TGCAGGCCCTT TCTGTGGTCT TCCTTTATCC AGGCTTTCTG CTCTCAGCCT TGCCTTCCTT				1559
TTTTCTGGGT CTCCTGGGCT GCCCCTTCA CTTGCAAAAT AAATCTTTGG TTTGTGCC				1619
TCTTCCCCAAA AAAAAAA				1636

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 466 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Val Arg Ser Val Ala Trp Ala Gly Phe Met Val Leu Leu Met Ile
-21 -20 -10

Pro Trp Gly Ser Ala Ala Lys Leu Val Cys Tyr Phe Thr Asn Trp Ala
-5 10

Gln Tyr Arg Gln Gly Glu Ala Arg Phe Leu Pro Lys Asp Leu Asp Pro
15 20 25

Ser Leu Cys Thr His Leu Ile Tyr Ala Phe Ala Gly Met Thr Asn His
30 35 40

Gln Leu Ser Thr Thr Glu Trp Asn Asp Glu Thr Leu Tyr Gln Glu Phe
45 50 55

Asn Gly Leu Lys Lys Met Asn Pro Lys Leu Lys Thr Leu Leu Ala Ile
60 65 70 75

Gly Gly Trp Asn Phe Gly Thr Gln Lys Phe Thr Asp Met Val Ala Thr
80 85 90

Ala Asn Asn Arg Gln Thr Phe Val Asn Ser Ala Ile Arg Phe Leu Arg
95 100 105

Lys Tyr Ser Phe Asp Gly Leu Asp Leu Asp Trp Glu Tyr Pro Gly Ser
110 115 120

Gln Gly Ser Pro Ala Val Asp Lys Glu Arg Phe Thr Thr Leu Val Gln
125 130 135

Asp Leu Ala Asn Ala Phe Gln Gln Glu Ala Gln Thr Ser Gly Lys Glu
140 145 150 155

Arg Leu Leu Leu Ser Ala Ala Val Pro Ala Gly Gln Thr Tyr Val Asp
160 165 170

Ala Gly Tyr Glu Val Asp Lys Ile Ala Gln Asn Leu Asp Phe Val Asn
175 180 185

Leu Met Ala Tyr Asp Phe His Gly Ser Trp Glu Lys Val Thr Gly His
190 195 200

Asn Ser Pro Leu Tyr Lys Arg Gln Glu Ser Gly Ala Ala Ala Ser
205 210 215

Leu Asn Val Asp Ala Ala Val Gln Gln Trp Leu Gln Lys Gly Thr Pro
220 225 230 235

Ala Ser Lys Leu Ile Leu Gly Met Pro Thr Tyr Gly Arg Ser Phe Thr
240 245 250

Leu Ala Ser Ser Ser Asp Thr Arg Val Gly Ala Pro Ala Thr Gly Ser
255 260 265

Gly Thr Pro Gly Pro Phe Thr Lys Glu Gly Gly Met Leu Ala Tyr Tyr
270 275 280

Glu Val Cys Ser Trp Lys Gly Ala Thr Lys Gln Arg Ile Gln Asp Gln
285 290 295

Lys Val Pro Tyr Ile Phe Arg Asp Asn Gln Trp Val Gly Phe Asp Asp

300	305	310	315
Val Glu Ser Phe Lys Thr Lys Val Ser Tyr Leu Lys Gln Lys Gly Leu			
320	325	330	
Gly Gly Ala Met Val Trp Ala Leu Asp Leu Asp Asp Phe Ala Gly Phe			
335	340	345	
Ser Cys Asn Gln Gly Arg Tyr Pro Leu Ile Gln Thr Leu Arg Gln Glu			
350	355	360	
Leu Ser Leu Pro Tyr Leu Pro Ser Gly Thr Pro Glu Leu Glu Val Pro			
365	370	375	
Lys Pro Gly Gln Pro Ser Glu Pro Glu His Gly Pro Ser Pro Gly Gln			
380	385	390	395
Asp Thr Phe Cys Gln Gly Lys Ala Asp Gly Leu Tyr Pro Asn Pro Arg			
400	405		410
Glu Arg Ser Ser Phe Tyr Ser Cys Ala Ala Gly Arg Leu Phe Gln Gln			
415	420	425	
Ser Cys Pro Thr Gly Leu Val Phe Ser Asn Ser Cys Lys Cys Cys Thr			
430	435	440	
Trp Asn			
445			

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1656 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 27..1424
 - (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) LOCATION: 90..1424

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

GCTGCAGCCT  GCCGCTGAGC  TGCATC  ATG  GTG  CGG  TCT  GTG  GCC  TGG  GCA  GGT      53
          Met  Val  Arg  Ser  Val  Ala  Trp  Ala  Gly
          -21   -20                                -15

```

TTC ATG GTC CTG CTG ATG ATC CCA TGG GGC TCT GCT GCA AAA CTG GTC
 Phe Met Val Leu Leu Met Ile Pro Trp Gly Ser Ala Ala Lys Leu Val
 -10 -5 1

TGC	TAC	TTC	ACC	AAC	TGG	GCC	CAG	TAC	AGA	CAG	GGG	GAG	GCT	CGC	TTC	149
Cys	Tyr	Phe	Thr	Asn	Trp	Ala	Gln	Tyr	Arg	Gln	Gly	Glu	Ala	Arg	Phe	
5					10					15					20	

CTG CCC AAG GAC TTG GAC CCC AGC CTT TGC ACC CAC CTC ATC TAC GCC
 Leu Pro Lys Asp Leu Asp Pro Ser Leu Cys Thr His Leu Ile Tyr Ala 197

DRAFT SEQUENCE

-40-

TTC GCT GGC ATG ACC AAC CAC CAG CTG AGC ACC ACT GAG TGG AAT GAC Phe Ala Gly Met Thr Asn His Gln Leu Ser Thr Thr Glu Trp Asn Asp 40 45 50	245
GAG ACT CTC TAC CAG GAG TTC AAT GGC CTG AAG AAG ATG AAT CCC AAG Glu Thr Leu Tyr Gin Glu Phe Asn Gly Leu Lys Lys Met Asn Pro Lys 55 60 65	293
CTG AAG ACC CTG TTA GCC ATC GGA GGC TGG AAT TTC AGC ACT CAG AAG Leu Lys Thr Leu Leu Ala Ile Gly Gly Trp Asn Phe Ser Thr Gln Lys 70 75 80	341
TTC ACA GAT ATG GTA GCC ACG GCC AAC AAC CGT CAG ACC TTT GTC AAC Phe Thr Asp Met Val Ala Thr Ala Asn Asn Arg Gln Thr Phe Val Asn 85 90 95 100	389
TCG GCC ATC AGG TTT CTG CGC AAA TAC AGC TTT GAC GGC CTT GAC CTT Ser Ala Ile Arg Phe Leu Arg Lys Tyr Ser Phe Asp Gly Leu Asp Leu 105 110 115	437
GAC TGG GAG TAC CCA GGA AGC CAG GGG AGC CCT GCC GTA GAC AAG GAG Asp Trp Glu Tyr Pro Gly Ser Gln Gly Ser Pro Ala Val Asp Lys Glu 120 125 130	485
CGC TTC ACA ACC CTG GTA CAG GAC TTG GCC AAT GCC TTC CAG CAG GAA Arg Phe Thr Thr Leu Val Gln Asp Leu Ala Asn Ala Phe Gln Gln Glu 135 140 145	533
GCC CAG ACC TCA GGG AAG GAA CGC CTT CTT CTG AGT GCA GCG GTT CCA Ala Gln Thr Ser Gly Lys Glu Arg Leu Leu Ser Ala Ala Val Pro 150 155 160	581
GCT GGG CAG ACC TAT GTG GAT GCT GGA TAC GAG GTG GAC AAA ATC GCC Ala Gly Gln Thr Tyr Val Asp Ala Gly Tyr Glu Val Asp Lys Ile Ala 165 170 175 180	629
CAG AAC CTG GAT TTT GTC AAC CTT ATG GCC TAC GAC TTC CAT GGC TCT Gln Asn Leu Asp Phe Val Asn Leu Met Ala Tyr Asp Phe His Gly Ser 185 190 195	677
TGG GAG AAG GTC ACG GGA CAT AAC AGC CCC CTC TAC AAG AGG CAA GAA Trp Glu Lys Val Thr Gly His Asn Ser Pro Leu Tyr Lys Arg Gln Glu 200 205 210	725
GAG AGT GGT GCA GCA GCC AGC CTC AAC GTG GAT GCT GCT GTG CAA CAG Glu Ser Gly Ala Ala Ala Ser Leu Asn Val Asp Ala Ala Val Gln Gln 215 220 225	773
TGG CTG CAG AAG GGG ACC CCT GCC AGC AAG CTG ATC CTT GGC ATG CCT Trp Leu Gln Lys Gly Thr Pro Ala Ser Lys Leu Ile Leu Gly Met Pro 230 235 240	821
ACC TAC GGA CGC TCC TTC ACA CTG GCC TCC TCA TCA GAC ACC AGA GTG Thr Tyr Gly Arg Ser Phe Thr Leu Ala Ser Ser Ser Asp Thr Arg Val 245 250 255 260	869
GGG GCC CCA GCC ACA GGG TCT GGC ACT CCA GGC CCC TTC ACC AAG GAA Gly Ala Pro Ala Thr Gly Ser Gly Thr Pro Gly Pro Phe Thr Lys Glu 265 270 275	917
GGA GGG ATG CTG GCC TAC TAT GAA GTC TGC TCC TGG AAG GGG GCC ACC Gly Gly Met Leu Ala Tyr Tyr Glu Val Cys Ser Trp Lys Gly Ala Thr 280 285 290	965
AAA CAG AGA ATC CAG GAT CAG AAG GTG CCC TAC ATC TTC CGG GAC AAC	1013

Lys Gln Arg Ile Gln Asp Gln Lys Val Pro Tyr Ile Phe Arg Asp Asn		
295	300	305
CAG TGG GTG GGC TTT GAT GAT GTG GAG AGC TTC AAA ACC AAG GTC AGC		1061
Gln Trp Val Gly Phe Asp Val Glu Ser Phe Lys Thr Lys Val Ser		
310	315	320
TAT CTG AAG CAG AAG GGA CTG GGC GGG GCC ATG GTC TGG GCA CTG GAC		1109
Tyr Leu Lys Gln Lys Gly Leu Gly Ala Met Val Trp Ala Leu Asp		
325	330	335
340		
TTA GAT GAC TTT GCC GGC TTC TCC TGC AAC CAG GGC CGA TAC CCC CTC		1157
Leu Asp Asp Phe Ala Gly Phe Ser Cys Asn Gln Gly Arg Tyr Pro Leu		
345	350	355
ATC CAG ACG CTA CGG CAG GAA CTG AGT CTT CCA TAC TTG CCT TCA GGC		1205
Ile Gln Thr Leu Arg Gln Glu Leu Ser Leu Pro Tyr Leu Pro Ser Gly		
360	365	370
ACC CCA GAG CTT GAA GTT CCA AAA CCA GGT CAG CCC TCT GAA CCT GAG		1253
Thr Pro Glu Leu Glu Val Pro Lys Pro Gly Gln Pro Ser Glu Pro Glu		
375	380	385
CAT GGC CCC AGC CCT GGA CAA GAC ACG TTC TGC CAG GGC AAA GCT GAT		1301
His Gly Pro Ser Pro Gly Gln Asp Thr Phe Cys Gln Gly Lys Ala Asp		
390	395	400
GGG CTC TAT CCC AAT CCT CGG GAA CGG TCC AGC TTC TAC AGC TGT GCA		1349
Gly Leu Tyr Pro Asn Pro Arg Glu Arg Ser Ser Phe Tyr Ser Cys Ala		
405	410	415
420		
GCG GGG CGG CTG TTC CAG CAA AGC TGC CCG ACA GGC CTG GTG TTC AGC		1397
Ala Gly Arg Leu Phe Gln Gln Ser Cys Pro Thr Gly Leu Val Phe Ser		
425	430	435
AAC TCC TGC AAA TGC TGC ACC TGG AAT TGAGTCGCTA AAGCCCCCTCC		1444
Asn Ser Cys Lys Cys Cys Thr Trp Asn		
440	445	
AGTCCCAGCT TTGAGGCTGG GCCCAGGATC ACTCTACAGC CTGCCCTCTG GGTTTCCCT		1504
GGGGGCCGCA ATCTGGCTCC TGCAAGGCCTT TCTGTGGTCT TCCTTTATCC AGGCTTTCTG		1564
CTCTCAGCCT TGCCTTCCTT TTTCTGGGT CTCCTGGGCT GCCCCTTCA CTTGCAAAAT		1624
AAATCTTGG TTTGTGCCCT TCAAAAAAA AA		1656

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 466 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Arg Ser Val Ala Trp Ala Gly Phe Met Val Leu Leu Met Ile
-21 -20 -15 -10

Pro Trp Gly Ser Ala Ala Lys Leu Val Cys Tyr Phe Thr Asn Trp Ala
-5 1 5 10

Gln Tyr Arg Gln Gly Glu Ala Arg Phe Leu Pro Lys Asp Leu Asp Pro

15 20 25

Ser Leu Cys Thr His Leu Ile Tyr Ala Phe Ala Gly Met Thr Asn His
30 35 40

Gln Leu Ser Thr Thr Glu Trp Asn Asp Glu Thr Leu Tyr Gln Glu Phe
45 50 55

Asn Gly Leu Lys Lys Met Asn Pro Lys Leu Lys Thr Leu Leu Ala Ile
60 65 70 75

Gly Gly Trp Asn Phe Ser Thr Gln Lys Phe Thr Asp Met Val Ala Thr
80 85 90

Ala Asn Asn Arg Gln Thr Phe Val Asn Ser Ala Ile Arg Phe Leu Arg
95 100 105

Lys Tyr Ser Phe Asp Gly Leu Asp Leu Asp Trp Glu Tyr Pro Gly Ser
110 115 120

Gln Gly Ser Pro Ala Val Asp Lys Glu Arg Phe Thr Thr Leu Val Gln
125 130 135

Asp Leu Ala Asn Ala Phe Gln Gln Glu Ala Gln Thr Ser Gly Lys Glu
140 145 150 155

Arg Leu Leu Leu Ser Ala Ala Val Pro Ala Gly Gln Thr Tyr Val Asp
160 165 170

Ala Gly Tyr Glu Val Asp Lys Ile Ala Gln Asn Leu Asp Phe Val Asn
175 180 185

Leu Met Ala Tyr Asp Phe His Gly Ser Trp Glu Lys Val Thr Gly His
190 195 200

Asn Ser Pro Leu Tyr Lys Arg Gln Glu Glu Ser Gly Ala Ala Ala Ser
205 210 215

Leu Asn Val Asp Ala Ala Val Gln Gln Trp Leu Gln Lys Gly Thr Pro
220 225 230 235

Ala Ser Lys Leu Ile Leu Gly Met Pro Thr Tyr Gly Arg Ser Phe Thr
240 245 250

Leu Ala Ser Ser Asp Thr Arg Val Gly Ala Pro Ala Thr Gly Ser
255 260 265

Gly Thr Pro Gly Pro Phe Thr Lys Glu Gly Gly Met Leu Ala Tyr Tyr
270 275 280

Glu Val Cys Ser Trp Lys Gly Ala Thr Lys Gln Arg Ile Gln Asp Gln
285 290 295

Lys Val Pro Tyr Ile Phe Arg Asp Asn Gln Trp Val Gly Phe Asp Asp
300 305 310 315

Val Glu Ser Phe Lys Thr Lys Val Ser Tyr Leu Lys Gln Lys Gly Leu
320 325 330

Gly Gly Ala Met Val Trp Ala Leu Asp Leu Asp Asp Phe Ala Gly Phe
335 340 345

Ser Cys Asn Gln Gly Arg Tyr Pro Leu Ile Gln Thr Leu Arg Gln Glu
350 355 360

DRAFT - 09/28/91

-43-

Leu Ser Leu Pro Tyr Leu Pro Ser Gly Thr Pro Glu Leu Glu Val Pro
365 370 375

Lys Pro Gly Gln Pro Ser Glu Pro Glu His Gly Pro Ser Pro Gly Gln
380 385 390 395

Asp Thr Phe Cys Gln Gly Lys Ala Asp Gly Leu Tyr Pro Asn Pro Arg
400 405 410

Glu Arg Ser Ser Phe Tyr Ser Cys Ala Ala Gly Arg Leu Phe Gln Gln
415 420 425

Ser Cys Pro Thr Gly Leu Val Phe Ser Asn Ser Cys Lys Cys Cys Thr
430 435 440

Trp Asn
445

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GACACTATAG AATAGGGC

18

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGGGATCATC AGCAGGACCA TGAAACCTGC CCAGGCCACA GACCGCACCA T

51

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TACATCTAGA ATTATGGCAA AACTGGTCTG CTACTTCACC

40

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid

02553510 - 02553511

-44-

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGATCTAACCTTAGGTGCCTGAAGACAAGT ATGG

34

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TACAGAAATTCTTATTTCACATCCGGCCCTG

29

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TACATCTAGACTCCATCCAGAAAAACAGGT ATGG

34

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCTAGAGTCGACCTGCAGGCATGCAAGCTT

30

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 50 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGCAAGCTTGAGAGCTCCGT TCCGCCACATGGTGCAGT GTGGCCTGGG

50

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAECTCTAGAC TAGGTGCCTG AAGGCAAGTA TG

32

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 373 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ala Lys Leu Val Cys Tyr Phe Thr Asn Trp Ala Gln Tyr Arg Gln Gly
1 5 10 15

Glu Ala Arg Phe Leu Pro Lys Asp Leu Asp Pro Ser Leu Cys Thr His
20 25 30

Leu Ile Tyr Ala Phe Ala Gly Met Thr Asn His Gln Leu Ser Thr Thr
35 40 45

Glu Trp Asn Asp Glu Thr Leu Tyr Gln Glu Phe Asn Gly Leu Lys Lys
50 55 60

Met Asn Pro Lys Leu Lys Thr Leu Leu Ala Ile Gly Gly Trp Asn Phe
65 70 75 80

Gly Thr Gln Lys Phe Thr Asp Met Val Ala Thr Ala Asn Asn Arg Gln
85 90 95

Thr Phe Val Asn Ser Ala Ile Arg Phe Leu Arg Lys Tyr Ser Phe Asp
100 105 110

Gly Leu Asp Leu Asp Trp Glu Tyr Pro Gly Ser Gln Gly Ser Pro Ala
115 120 125

Val Asp Lys Glu Arg Phe Thr Thr Leu Val Gln Asp Leu Ala Asn Ala
130 135 140

Phe Gln Gln Glu Ala Gln Thr Ser Gly Lys Glu Arg Leu Leu Leu Ser
145 150 155 160

Ala Ala Val Pro Ala Gly Gln Thr Tyr Val Asp Ala Gly Tyr Glu Val
165 170 175

Asp Lys Ile Ala Gln Asn Leu Asp Phe Val Asn Leu Met Ala Tyr Asp
180 185 190

Phe His Gly Ser Trp Glu Lys Val Thr Gly His Asn Ser Pro Leu Tyr
195 200 205

00238618-0022896

0000000000000000

Lys Arg Gln Glu Glu Ser Gly Ala Ala Ala Ser Leu Asn Val Asp Ala
210 215 220
Ala Val Gln Gln Trp Leu Gln Lys Gly Thr Pro Ala Ser Lys Leu Ile
225 230 235 240
Leu Gly Met Pro Thr Tyr Gly Arg Ser Phe Thr Leu Ala Ser Ser Ser
245 250 255
Asp Thr Arg Val Gly Ala Pro Ala Thr Gly Ser Gly Thr Pro Gly Pro
260 265 270
Phe Thr Lys Glu Gly Gly Met Leu Ala Tyr Tyr Glu Val Cys Ser Trp
275 280 285
Lys Gly Ala Thr Lys Gln Arg Ile Gln Asp Gln Lys Val Pro Tyr Ile
290 295 300
Phe Arg Asp Asn Gln Trp Val Gly Phe Asp Asp Val Glu Ser Phe Lys
305 310 315 320
Thr Lys Val Ser Tyr Leu Lys Gln Lys Gly Leu Gly Gly Ala Met Val
325 330 335
Trp Ala Leu Asp Leu Asp Asp Phe Ala Gly Phe Ser Cys Asn Gln Gly
340 345 350
Arg Tyr Pro Leu Ile Gln Thr Leu Arg Gln Glu Leu Ser Leu Pro Tyr
355 360 365
Leu Pro Ser Gly Thr
370

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 373 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Lys Leu Val Cys Tyr Phe Thr Asn Trp Ala Gln Tyr Arg Gln Gly
1 5 10 15
Glu Ala Arg Phe Leu Pro Lys Asp Leu Asp Pro Ser Leu Cys Thr His
20 25 30
Leu Ile Tyr Ala Phe Ala Gly Met Thr Asn His Gln Leu Ser Thr Thr
35 40 45
Glu Trp Asn Asp Glu Thr Leu Tyr Gln Glu Phe Asn Gly Leu Lys Lys
50 55 60
Met Asn Pro Lys Leu Lys Thr Leu Leu Ala Ile Gly Gly Trp Asn Phe
65 70 75 80
Gly Thr Gln Lys Phe Thr Asp Met Val Ala Thr Ala Asn Asn Arg Gln
85 90 95
Thr Phe Val Asn Ser Ala Ile Arg Phe Leu Arg Lys Tyr Ser Phe Asp
100 105 110

Gly Leu Asp Leu Asp Trp Glu Tyr Pro Gly Ser Gln Gly Ser Pro Ala
115 120 125

Val Asp Lys Glu Arg Phe Thr Thr Leu Val Gln Asp Leu Ala Asn Ala
130 135 140

Phe Gln Gln Glu Ala Gln Thr Ser Gly Lys Glu Arg Leu Leu Leu Ser
145 150 155 160

Ala Ala Val Pro Ala Gly Gln Thr Tyr Val Asp Ala Gly Tyr Glu Val
165 170 175

Asp Lys Ile Ala Gln Asn Leu Asp Phe Val Asn Leu Met Ala Tyr Asp
180 185 190

Phe His Gly Ser Trp Glu Lys Val Thr Gly His Asn Ser Pro Leu Tyr
195 200 205

Lys Arg Gln Glu Glu Ser Gly Ala Ala Ala Ser Leu Asn Val Asp Ala
210 215 220

Ala Val Gln Gln Trp Leu Gln Lys Gly Thr Pro Ala Ser Lys Leu Ile
225 230 235 240

Leu Gly Met Pro Thr Tyr Gly Arg Ser Phe Thr Leu Ala Ser Ser Ser
245 250 255

Asp Thr Arg Val Gly Ala Pro Ala Thr Gly Ser Gly Thr Pro Gly Pro
260 265 270

Phe Thr Lys Glu Gly Gly Met Leu Ala Tyr Tyr Glu Val Cys Ser Trp
275 280 285

Lys Gly Ala Thr Lys Gln Arg Ile Gln Asp Gln Lys Val Pro Tyr Ile
290 295 300

Phe Arg Asp Asn Gln Trp Val Gly Phe Asp Asp Val Glu Ser Phe Lys
305 310 315 320

Thr Lys Val Ser Tyr Leu Lys Gln Lys Gly Leu Gly Gly Ala Met Val
325 330 335

Trp Ala Leu Asp Leu Asp Asp Phe Ala Gly Phe Ser Cys Asn Gln Gly
340 345 350

Arg Tyr Pro Leu Ile Gln Thr Leu Arg Gln Glu Leu Ser Leu Pro Tyr
355 360 365

Leu Ser Ser Gly Thr
370

085321 082096